A SIMULATION APPROACH TO EVALUATE COMBINING FORECASTS METHODS

by

HO KWONG-SHING LAWRENCE

何光誠

MBA PROJECT REPORT

Presented to

The Graduate School

In Partial Fulfilment

of the Requirements for the Degree of

MASTER OF BUSINESS ADMINISTRATION

TWO-YEAR MBA PROGRAMME

THE CHINESE UNIVERSITY OF HONG KONG

May 1994

APPROVAL

Name: Ho Kwong-shing Lawrence

Degree: Master of Business Administration

Title of Project: A Simulation Approach to Evaluate Combining Forecasts Methods

Dr Lau Kinnam

Date Approved: May 15, 94

ABSTRACT

The purpose of this study is to evaluate four different sales forecasting combination methods by using a Monte Carlo simulation approach. The methods to be analyzed are analytical method of optimum forecasts combination, ordinary least square (OLS) regression method, linear programming (LP) method as well as the boundary value method (BVM). For each experiment two hundred replicates are run with the number of individual sales forecasters set to be five. We use small sample size of five time periods as well as normal sample size of thirty time periods to see whether there is difference in the ranking of the four methods. Having taken different variance and covariance structures into consideration when combining the individual sales forecasts, it has been found that the OLS method ranks the first in terms of mean square error (MSE) while the LP method ranks the first in terms of mean absolute deviation (MAD) and sum of absolute percentage error (APE). They dominate their respective forecast error measures. In particular, with larger sample size the dominant positions of them are enhanced. The other two methods, namely, the analytical and BVM method lag behind to a significant extent. On the other hand, when sample size is increased, the forecasting power of the four methods become closer to each other.

TABLE OF CONTENTS

ACKNOWLEDGEMENT

The completion of this project cannot be possible without the sincere assistance and advice from my project advisor Dr. Lau Kinnam. I am obliged to express my heartily thanks to him. Moreover, his patience to discuss matters relating to the whole research process is highly appreciated.

In addition, I am grateful to Mr. Shi Jian Qing for his assistance in writing and editing the computer program so that the simulation process could be undertaken successfully.

Lastly, I would like to thank my wife who shares the pressure with me when we are pursuing advanced studies at the same time. This experience is by no means easy but should be highly rewarding from the standpoint of personal development.

CHAPTER I

INTRODUCTION

Marketing manager often faces the situation to forecast the sales volume or market share in short and medium terms in order to facilitate the planning and control purpose. Besides, an accurate sales forecast is vital to other functional activities within any business firms whether they are trading-oriented or production-oriented. The issue to combine sales forecasts of different salespersons becomes the practical problem for marketing manager to tackle. Next section will be a review of the forecasts combination literature in general.

Chapter II discusses the combining sales forecasts literature and its related problems in actual marketing situation.

Chapter III is about the experimental design of the project. Simulation is adopted to test the relative usefulness of four different methods in combining sales forecasts. Various testing statistics are calculated and the ranking of testing procedures is done to evaluate these methods.

Chapter IV presents the simulation findings under certain assumptions of simulated actual sales data, variance and covariance among different sales forecasters and time periods concerned.

The last chapter is summary and conclusion which summarises our findings of the Monte Carlo experiments of sales forecast combination.

LITERATURE REVIEW

As one of the important tools to facilitate business planning and decision making, forecasting techniques are widely used in both private organisations and public sector. With the popularity of computer and associated software development the forecasting process becomes easier to manage than before.

Regarding the forecasting methods currently in use, Makridakis and Wheelwright (1989) divide it into three major approaches. The first one is judgmental method which employs individual judgments or committee agreements or decisions to generate the forecast. The experience of using this method is that it addresses a vast majority of forecasting needs.

Next comes the quantitative method which is the focus of the majority of forecasting literature especially after

1960s when statistically sophisticated methods could be used in a more and more computerised environment. There are three subcategories under this method. Time-series method uses time as a reference to uncover any historical pattern and a time-based extrapolation of those pattern becomes the basis for forecasting. Explanatory method like regression tries to identify the causality of observed outcomes in the past and applies the found relationships to forecast future events. Monitoring method seeks to identify changes in the underlying patterns and relationships when extrapolation of them is not appropriate.

The third approach is the technological method which address long-term issue of a technological, societal, economic or political nature. Under this method it can be extrapolative, analogy-based, expert-based as well as normative-based. The well known example of this method is the Delphi method which uses expert opinions of a panel to deal with specific problems. Informed intuitive judgment of different experts figures out those forecasting and influencing factors not considered by other people or included in the calculation of statistical forecast (Bolt, 1988).

Forecast combination attracts attention and research interest of academicians after Bates and Granger (1969) proposed their theory about combining forecasts and tested several techniques for combining point forecasts. Further development of theoretical statistical models by Dickinson (1973) and Bunn (1975) enriched the literature by using different approaches to forecasts combination. Dickinson formulated a minimum-variance model while Bunn adopted a practical Bayesian approach to using multiple forecasts. The article by Granger and Ramanathan (1984) influences conventional forecast combination methods by pointing out that standard regression techniques were equivalent to constrained ordinary least square estimation procedure. They argued that instead of constraining the combining weights to sum to one and setting the intercept term to be zero, running unconstrained least squares could give a better fit to past data. However, setting constraints can improve the robustness of the combination in forecasting.

Although Granger and Ramanathan's article was not entirely original in the application of regression techniques and has been contested by other researchers theoretically and empirically (Clemen 1986; Mills and Stephenson 1985), it provided an important impetus for the use of sophisticated econometric method in doing combined forecast (Clemen 1989). On the other hand, to use the least-square regression analysis the multicollinearity problem should also be addressed in addition to the objective of minimising error variance of combined forecast.

Wilson and Keating (1990) suggested it might be unwise to select the most accurate forecasting method but to combine the forecasts already made. They emphasised on the independent information contained in different forecasts and forecast improvement by way of reduction of root mean squared error (RMSE) when forecast combination is done. They concluded that "in general a combined forecast will have a much smaller error, as measured by RMSE, unless individual forecasting models are almost equally good and their forecast errors are highly correlated." (p.338). On the other hand, Granger and Newbold (1973) pointed out the conditional efficiency issue so as to establish an objective standard to measure the benefit of combining individual forecasting models in terms of the incremental improvement in forecasting accuracy.

With a set of available individual forecasting models, Russell and Adam (1987) suggested that it was more appropriate to selectively include some of them rather than use all of them in the process of combination. Higher forecasting error as a result of information redundancy (this issue will be discussed in later part) may occur due to inclusion of the less effective or inaccurate models. Newbold, Zumvalt and Kannan (1987) indicated that if all the individual models were included in forecast combination, the assignment of negative weights to some constituent models followed naturally. But this phenomenon is not justified on grounds that it is contrary to the spirit of forecast combination. So the usual procedure to overcome negative weights is to drop those models with negative weights from the combination model and re-calculate the optimal weights of the remaining individual models.

CHAPTER II

COMBINING SALES FORECASTS

In the context of business forecasting, making sales forecasts is vital to planning and budgeting activity. Eby and O'Neill (1977) defined a sales forecast as "the sales volume a firm expects to realize during a designated future time period. It is a projection based upon a carefully formulated marketing plan, along with an evaluation of market factors which may have an influence on future sales". They pointed out the important applications of sales forecasts in connection with undertaking changes in plant capacity, acquiring the proper labour supply, production scheduling, inventory control, maintaining optimum product mixes, planning sales force activities and so forth.

In a recent article by Smith, McIntyre and Achabal (1994) they developed a two-stage sales forecasting procedure using discounted least squares. Their method is implemented in two stages in which the sum of squares of forecast errors are minimised. Standard regression analyses are performed in the process of estimation and updating. In stage one the coefficients of the controllable and environmental variables are estimated and then in stage two discounted least squares smoothing procedures that update selected key parameters are undertaken so that the impact on the sales forecast can be accounted for as marketing environment changes. The accurate specification and estimation of forecast models is essential to the forecast combination process because individual models accuracy influences the size of the forecast error of the combined forecasts.

Actually there are various kinds of techniques to make sales forecasts by means of estimating market demand for product or service. Of these the sales force combination method deserves attention. For this method every member of the sales force provides his or her individual sales forecast for the predetermined time horizon. These individual forecasts are then checked, discussed and combined at upper management level so that a combined forecast is arrived for planning, budgeting and decision-making purposes. Churchill et al. (1993) considered the primary advantage of the sales force combination method is that it uses specialised knowledge of salespersons who are closest to the market and therefore makes the final sales forecast fairly accurate. Moreover, it aids in directing the sales efforts because the formulation of sales quota which is based on the combined sales forecast involves the field salesperson directly. However, the disadvantage of the method is that the vested interest of individual salesperson leads to bias which is costly to correct. Or elaborate schemes are required to counteract bias.

To apply the techniques of sales forecast combination some practical problems arise when marketing manager seeks to obtain the accurate sales forecasts (Moriarty 1990).

First, small sample sizes of individual forecast models make it difficult for calibration of forecast combination model (Sessions and Chatterjee 1989). The number of observations is small compared to the combination model parameters. In a marketing context it is very often to incorporate new individual models to the existing combination model or to formulate a new combination model due to product development from which observations are lacking.

The second problem is about dependent variable instability. Sales volume is one of the examples of this property because the underlying pattern changes due to keen competition, advertising or new product activity (Weitz 1985). This implies that the weights or importance attached to individual forecasting models in the forecast combination changes also. Together with the phenomenon of non-stationarity of weights, the structural change in forecast errors pattern of different models reduces the benefit of forecast combination. Some methods of combination to cope with these problems like both adaptive estimation of covariance matrix of forecast errors of individual models and econometric methods for structural change are proposed by researchers. More recently, state-space method and Kalman filtering are sought to model non-stationary weights (Bunn 1989).

The third problem is redundancy of information sets upon which alternative individual forecasting models are built (Clemen and Winkler 1985, Granger 1989). This property occurs when the individual forecasting models that are included in the combination model contain similar or same set of explanatory variables. This situation is not uncommon in the marketing situations. As a result, weighted combination of individual models brings about marginal improvement in the forecasting performance of the forecast combination (Newbold, Zumwalt and Kannan 1987).

Moriarty pointed out that with the properties of small sample sizes, instability in the variable to be forecast and constituent model redundancy, two implementation problems have to be resolved if forecast combination is performed. The first problem is the specification and estimation of the respective linear weights which represent the relative accuracy of individual models and the correlations of their disturbance terms. The second problem is about the use justification of forecast combination considering that the incremental value of an additional individual model to a combination may not warrant maintaining it in the company's marketing forecasting system.

He demonstrated that the analytical model of optimum forecast combination could not resolve the aforementioned three impediments in one situation or the others. In response to this issue, he proposed a *boundary value model* which he defined as "a forecast combination model whose individual weights are chosen such that they approximate the interrelationship among the individual forecast models' accuracy values and disturbance term correlation while assuming values at a boundary on the range of weights". Examples of weights he has suggested are equal weights for all n individual models ($w_i = 1/n$), equal weights for a subset (= k) of models ($w_i = 1/k$; $k = 2,\ldots,n-1$) or all weight placed to only one model (the single best model in which $w_1 = 1$ and $w_i = 0$ for $i = 2,\ldots,n$). In dealing with small sample size of prior forecasts property, selection of a BVM based on a marketing forecaster's limited judgment as well as insufficient empirical data may be preferable to the analytical method especially if some information about accuracy of individual models and errors correlations is known. On the other hand, a fixed-weight BVM as a alternative choice could compensate for the dependent variable instability property, i.e., volatility in the accuracy of individual models, intercorrelations of errors and the resulting changes in optimal weights. He believed that a equal-weights BVM would be equally competitive with a more elaborate bivariate ARCH model which accounts for process instability when judging from the magnitude of forecast mean squared error. Lastly, under the circumstances that different models use similar source of information, individual model redundancy creates an estimation problem that combination weights may vary erratically. Little forecast error reduction or negative weights are encountered even though estimation can be accomplished. So he recommended the BVMs to resolve the information redundancy problem.

Regarding the choice of different forms of BVM, Moriarty suggested to use a single best BVM when one model dominates the others in forecast accuracy and no incremental information can be obtained from error structure. In fact, this choice is the same as the concept of "encompassing" when one model outperforms the others and encompass them in forecasting. The BVM with a subset of equal weights as mentioned before is appropriate in cases that the k (<n) individual models have competitive accuracy and the other forecast models are either not competitively accurate or have high positive error intercorrelations with the k models. The equal-weights BVM has been proposed as an alternative when firms cannot determine the weights attached to individual forecast models in advance. This model is only the practical one to choose when no sample forecasts are available for calibration of combination weights. This method is the same as taking a simple average of all the individual forecast models and may not

give efficiency gain on forecast error reduction in some cases. But its major advantage is that it avoids the risk of making a priori choice of a best individual model whose forecast error is in fact much larger than expected. With more information accumulates as sample size grows larger when time passes, refined BVMs can be tested and adopted to reflect the changes in marketing environment and underlying patterns of dependent variables.

Notwithstanding the advantages that Moriarty's BVMs bring to forecast combination in a number of marketing situations that the application of the analytical method of optimum forecast combination is difficult if not impossible, two important questions are still left unanswered, i.e., how to filter out those individual forecasters who should not included in the combination process for the purpose of improving forecasting accuracy; and how to determine the optimal weights of the constituent individual models within the combination model.

CHAPTER III

EXPERIMENTAL DESIGN

This study employs the simulation techniques to generate sales forecast data from individual forecasting units. In simulation, computer programs and models are built up to imitate the real life situations and then statistical data are collected for further analysis. Since experiments can be done for many times by using computer, it facilitates the process of data collection as well as quantitative analyses.

The number of replicates for each simulation is set to be two hundred. In order to arrive at forecasts combination using different methods, five sales forecasters ($n = 5$) are presumed. The number of time periods t (sample sizes) are five and thirty respectively so as to compare the small sample size case with the normal sample size case.

Initially it is assumed that the mean of sale volume is 1,000 with standard deviation of 100. The carrying out of different simulations depends on the assumption of the population variance-covariance matrix. The variances of individual forecast consist of three different forms: 1) all of them have equal variance of 10,000; 2) the variance is in descending order so that the five individual variances are 10,000; 8,100; 6,400; 4,900; 3,600 respectively; 3) two of them have equal variance of 10,000 while the other three have equal variance of 6,400. With respect to the covariance structure, seven cases are considered for simulation purpose. They are as follows:

1) no correlation	:	p=0
2) low positive correlation	:	p=0.2
3) median positive correlation	:	p=0.5
4) high positive correlation	:	p=0.8
5) low negative correlation	:	p=-0.2
6) median negative correlation	:	p=-0.5
7) high negative correlation	:	p=-0.8

where p is the correlation coefficient between two different forecasters. Hence for two different time periods (sample sizes) there are totally forty two sets of simulation results for study.

Individual sales forecasts are assumed to be normally distributed. But the actual sales figures at time t have different kinds of distribution which are classified as normal distribution, Laplace distribution, Cauchy distribution as well as uniform distribution. The purpose of using different assumptions of actual sales distribution is to find out the relative usefulness of the four different estimation methods in particular situations. Further, individual forecasts are assumed to be unbiased estimator of the actual sales volume figures.

Then the testing procedure in the Monte Carlo experiment is discussed. Four linear weights estimation methods are run and compared. They consist of :

1) Analytical (ANA) method to minimise the forecasts combination variance. The model is specified as follows:

$$F_{*t} = F_{it} + \epsilon_{it} \qquad i = 1, \ldots\ldots, n.$$

and $\epsilon_{it} \sim N(0, \sigma_i^2)$:

where F_{*t} is actual sales, F_{it} is estimated sales and ϵ_{it} is the disturbance term of the i^{th} forecaster in time period t.

The combination forecasting model is given by:

$$F_{ct} = \Sigma^n{}_{i=1}\ \alpha_i F_{it}$$

and $\Sigma^n{}_{i=1}\ \alpha_i = 1$:

where F_{ct} is the combination forecast of F_t and α_i is the forecast weight for the i^{th} forecast. The combination forecast variance is related to the individual forecaster' variance and disturbance term correlation p_{ij} by the equation of :

$$\sigma_c^2 = \alpha' \Sigma \alpha$$

where $\alpha' = [\ \alpha_1\ \alpha_2\ \ldots\ \alpha_n]$ and the variance-covariance matrix (Σ) is

$$\Sigma = \begin{bmatrix} \sigma_1^2 & p_{12}\sigma_1\sigma_2 & \cdots & p_{1n}\sigma_1\sigma_n \\ p_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & p_{2n}\sigma_2\sigma_n \\ \vdots & \vdots & & \vdots \\ p_{1n}\sigma_1\sigma_n & p_{2n}\sigma_2\sigma_n & \cdots & \sigma_n^2 \end{bmatrix}$$

The optimal minimum variance α_i forecast weights are calculated as :

$$\alpha'_* = u'\Sigma^{-1} \ /\ (u'\Sigma^{-1}u)$$

where u is a n X 1 unity vector and the asterisk subscript represents the optimality condition.

We shall use the above model to simulate the results obtained under this analytical method.

2) Ordinary least square (OLS) regression method

$F^*_t = w_1.F_{1t} + w_2.F_{2t} + \ldots\ldots + w_n.F_{nt} + \epsilon_t$

where $\Sigma w_i = 1$ for all $w_i >= 0$ and $t = 1, \ldots, T$

The OLS method is to minimise the square of forecast errors for the whole time period.

3) Linear programming (LP) method to obtain least absolute deviation

Objective function : min $\Sigma (e^+_t + e^-_t)$

subject to

$F^*_t = w_1.F_{1t} + w_2.F_{2t} + \ldots\ldots + w_n.F_{nt} + e^+_t + e^-_t$

$\Sigma w_i = 1$

$w_i >= 0$ for $t = 1, \ldots, T$

4) Boundary value model (BVM) method

As mentioned in the last chapter there are three versions of the BVM. We choose the 'equal weight' version so that the weight for every forecaster is simply the average of the number of forecasters. In the present case it is 0.2 since five forecasters are selected. The equation is given by :

$$F^*_t = (F_{it})/n + \epsilon_t$$

$$= 0.2F_{it} + \epsilon_t$$

For each replicate, average value of the combined forecast estimate and variance of the actual sales are calculated. In order to evaluate the above four different approaches to forecast combination, the three frequently encountered forecast error measures are computed also. They are mean absolute deviation (MAD), mean squared error (MSE) and sum of absolute percentage error (APE).

In each experiment, grand statistics of the above five items are computed. This is simply dividing the relevant statistics by the number of replicates which is two hundred in this case.

Ranking of procedures is undertaken so as to see the number of times that each method win or lose. Because there are four methods, it follows that the rank is from one to four. If a certain method ranks the first under certain conditions for most of the replicates and/or has the lowest value of a particular forecast error measure, it can be viewed as the winner and be considered for use in forecast combination.

CHAPTER IV

SIMULATION RESULTS

This chapter is about the simulation findings for different cases under assumption of different population variance-covariance matrix. Totally there are forty two cases to be analyzed. Small sample size situation contains twenty one cases and normal sample size situation contains another twenty one cases. The Appendix contains the ranking of the four procedures by showing the number of times each method wins under different distribution of the actual sales figures. The following paragraphs will discuss some cases of grand statistics of forecast error measures under different distribution of actual sales.

(A) Small Sample Size (T=5)

1) Equal Variance and p=0

Normal Distribution

Method	MSE	MAD	APE
ANA	11212	83	0.43
OLS	6195	59	0.30
LP	7234	54	0.28
BVM	11726	85	0.44

Cauchy Distribution

Method	MSE	MAD	APE
ANA	3571165	427	2.78
OLS	3504683	406	2.57
LP	3536408	394	2.48
BVM	3573122	429	2.70

Uniform Distribution

Method	MSE	MAD	APE
ANA	11563	91	0.46
OLS	6922	67	0.34
LP	7905	62	0.31
BVM	12124	93	0.47

Laplace Distribution

Method	MSE	MAD	APE
ANA	10852	78	0.40
OLS	9357	59	0.31
LP	9554	54	0.28
BVM	10360	81	0.42

It has been found that OLS has the smallest MSE for all the four kinds of distribution while LP has the smallest MAD and APE for all the four kinds of distribution. Both ANA and BVM have similar size of forecast errors with the former one slightly better.

2) Variance in descending order and p=0.2

Normal Distribution

Method	MSE	MAD	APE
ANA	11139	83	0.42
OLS	7262	65	0.33
LP	8174	61	0.31
BVM	12144	88	0.44

Cauchy Distribution

Method	MSE	MAD	APE
ANA	3569957	426	2.69
OLS	3517301	411	2.57
LP	3544340	402	2.50
BVM	3569455	429	2.66

Uniform Distribution

Method	MSE	MAD	APE
ANA	11548	91	0.46
OLS	7963	72	0.36
LP	8740	68	0.34
BVM	12519	94	0.47

Laplace Distribution

Method	MSE	MAD	APE
ANA	11427	78	0.40
OLS	8310	65	0.33
LP	9357	60	0.31
BVM	12512	82	0.43

Again, it has been found that OLS has the smallest MSE for all the four kinds of distribution while LP has the smallest MAD and APE for all the four kinds of distribution. Both ANA and BVM have similar size of forecast errors with the former one slightly better.

(B) Normal Sample Size (T=30)

1) Equal Variance and p=0

Normal Distribution

Method	MSE	MAD	APE
ANA	11604	86	2.625
OLS	10078	81	2.481
LP	10994	79	2.409
BVM	11867	86	2.653

Cauchy Distribution

Method	MSE	MAD	APE
ANA	35453687	567	7.58
OLS	35356874	571	8.46
LP	35424693	559	7.66
BVM	35451671	568	7.60

Uniform Distribution

Method	MSE	MAD	APE
ANA	11559	91	2.76
OLS	10375	85	2.58
LP	10964	82	2.51
BVM	11867	91	2.78

Laplace Distribution

Method	MSE	MAD	APE
ANA	11634	79	2.45
OLS	10440	76	2.35
LP	11068	74	2.27
BVM	11960	81	2.49

Similar to the small sample size case it has been found that OLS has the smallest MSE for all the four kinds of distribution while LP has the smallest MAD and APE for all the four kinds of distribution. Both ANA and BVM have similar size of forecast errors with the former one slightly better. When the sample size has been increased, the difference between the forecast errors among the four methods shows signs of decrease. It implies that the forecasting power of them will be close to each other when the sample size increases. This results matches the prediction of theory.

2) Variance in descending order and p=0.2

Normal Distribution

Method	MSE	MAD	APE
ANA	11608	86	2.625
OLS	10626	82	2.506
LP	11082	80	2.449
BVM	12162	88	2.688

Cauchy Distribution

Method	MSE	MAD	APE
ANA	35455776	567	7.63
OLS	35390559	569	8.19
LP	35430528	560	7.63
BVM	35458051	569	7.59

Uniform Distribution

Method	MSE	MAD	APE
ANA	11588	91	2.76
OLS	10619	86	2.61
LP	11079	84	2.56
BVM	12198	92	2.82

Laplace Distribution

Method	MSE	MAD	APE
ANA	11655	79	2.45
OLS	10686	77	2.37
LP	11159	75	2.31
BVM	12323	82	2.54

Again, it has been found that OLS has the smallest MSE for all the four kinds of distribution while LP has the smallest MAD and APE for all the four kinds of distribution. Both ANA and BVM have similar size of forecast errors with the former one slightly better. Moreover, a larger sample size of thirty makes the forecasting power of the four different methods to be close to each other.

CHAPTER V

SUMMARY AND CONCLUSION

The findings of the simulation results are summarised and listed in the following :

(A) Small Sample Size Case (T=5)

1) The ordinary least square method (OLS) dominates other methods in terms of the MSE measure with linear programming (LP) method ranks the second. This phenomenon appears in all four different kinds of distribution of actual sales data, i.e., normal, Cauchy, uniform and Laplace and different covariance structures among individual forecasters;

2) The linear programming (LP) method dominates other methods in terms of both the MAD and APE measures with ordinary least square (OLS) method ranks the second. This phenomenon appears in all four different kinds of distribution of actual sales data, i.e., normal, Cauchy, uniform and Laplace and different covariance structures among individual forecasters;

3) In case individual forecasters are more positively correlated (from 0.2 to 0.8), the dominant positions of OLS and LP in aforementioned error measures shows signs of reduction. The number of replicates they win

is decreased by about 5%;

4) In case individual forecasters are more negatively correlated (from -0.2 to -0.8), the respective dominant positions of OLS and LP are not affected;

5) Analytical (ANA) method and boundary value method (BVM) occupies the ranks of third and forth for most of the replicates. ANA is more accurate than the BVM marginally.

(B) Normal Sample Size Case (T=30)

1) When sample size is increased to thirty, the dominant position of OLS in MSE measure and LP in MAD and APE measures are enhanced a lot. In particular, they win nearly all the two hundred replicates in the respective error measures. This applies to different distribution of simulated actual sales data as well as different covariance structure among individual sales forecasters;

2) In Cauchy distribution, although LP ranks the first in APE measure the number of winning the 200 replicates is around 150. Owing to the special nature of this type of distribution (whose central tendency are location and sigma instead of mean and standard deviation), LP's dominant status is not as strong as in other cases;

3) Similar to the small sample size situation, both the analytical and BVM methods fall behind the OLS and LP methods;

4) The magnitude of sample size does affect the forecasting power of the four methods. With larger sample size their forecast errors show smaller differences and this corresponds to the results predicted from theory.

Since OLS is to minimise the squared error, it follows that it works best when MSE is used. On the other hand, since LP is to minimise the absolute deviation from the actual sales value, it follows that it works best when MAD and APE are used. As the analytical method is to minimise the variance of the forecast combination by estimating the relevant linear weights, it is not as good as the OLS and LP methods when those forecast error measures are considered. The BVM (equal weights for the five individual forecasters) is not able to achieve higher ranks because of the uncertainty and complexity of the forecasting structure. Perhaps other forms of BVM can be tested to see its robustness.

(C) Limitation of the Study

The above simulation exercise is based on the assumption that every sales forecast are unbiased so that its expected value is the same as the actual sales volume. Future research can be done to include the biased forecast situation so as to give us more insights into the choice of combined forecast method in different situations.

Moreover, a larger number of individual forecasters can be performed (say 10 or 30 forecasters) in the simulation process and more complex covariance structures among them are considered, e.g., positive and negative correlations co-exit in the pool of forecasters.

APPENDIX

THE RANKING OF PROCEDURES UNDER THE SITUATIONS OF SMALL AND NORMAL SAMPLE SIZE AND DIFFERENT CORRELATION COEFFICIENTS

(A) Small Sample Size (T=5)

Case 1 : Equal Variance, p=0

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (197)*	LP (194)	LP (187)
Cauchy	OLS (192)	LP (186)	LP (144)
Uniform	OLS (198)	LP (196)	LP (192)
Laplace	OLS (196)	LP (197)	LP (189)

* The number of times the winning method ranks the first in the two hundred replicates. This applies to following analyses.

Case 2 : Descending Variance, p=0

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (198)	LP (193)	LP (186)
Cauchy	OLS (194)	LP (178)	LP (138)
Uniform	OLS (196)	LP (188)	LP (183)
Laplace	OLS (199)	LP (195)	LP (188)

Case 3 : Variance of two greater than that of other three, p=0

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (197)	LP (192)	LP (188)
Cauchy	OLS (193)	LP (181)	LP (145)
Uniform	OLS (199)	LP (191)	LP (185)
Laplace	OLS (200)	LP (194)	LP (188)

Case 4 : Equal Variance, p=0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (191)	LP (191)	LP (185)
Cauchy	OLS (186)	LP (176)	LP (139)
Uniform	OLS (198)	LP (191)	LP (181)
Laplace	OLS (197)	LP (192)	LP (185)

Case 5 : Descending Variance, p=0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (196)	LP (184)	LP (175)
Cauchy	OLS (194)	LP (173)	LP (135)
Uniform	OLS (194)	LP (190)	LP (184)
Laplace	OLS (198)	LP (190)	LP (181)

Case 6 : Variance of two greater than that of other three , p=0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (192)	LP (190)	LP (183)
Cauchy	OLS (189)	LP (178)	LP (143)
Uniform	OLS (195)	LP (189)	LP (180)
Laplace	OLS (198)	LP (192)	LP (187)

Case 7 : Equal Variance, p=0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (186)	LP (180)	LP (175)
Cauchy	OLS (181)	LP (167)	LP (131)
Uniform	OLS (193)	LP (187)	LP (179)
Laplace	OLS (195)	LP (193)	LP (187)

Case 8 : Descending Variance, p=0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (182)	LP (174)	LP (166)
Cauchy	OLS (184)	LP (164)	LP (133)
Uniform	OLS (189)	LP (179)	LP (168)
Laplace	OLS (193)	LP (185)	LP (176)

Case 9 : Variance of two greater than that of other three , p=0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (188)	LP (179)	LP (173)
Cauchy	OLS (181)	LP (166)	LP (130)
Uniform	OLS (190)	LP (185)	LP (176)
Laplace	OLS (192)	LP (189)	LP (184)

Case 10 : Equal Variance, p=0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (178)	LP (166)	LP (155)
Cauchy	OLS (170)	LP (149)	LP (119)
Uniform	OLS (179)	LP (161)	LP (155)
Laplace	OLS (171)	LP (173)	LP (167)

Case 11 : Descending Variance, p=0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (160)	LP (157)	LP (156)
Cauchy	OLS (170)	LP (160)	LP (126)
Uniform	OLS (164)	LP (157)	LP (149)
Laplace	OLS (172)	LP (167)	LP (162)

Case 12 : Variance of two greater than that of other three , p=0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (184)	LP (171)	LP (164)
Cauchy	OLS (185)	LP (166)	LP (127)
Uniform	OLS (189)	LP (178)	LP (168)
Laplace	OLS (184)	LP (185)	LP (179)

Case 13 : Equal Variance, p=-0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (199)	LP (198)	LP (196)
Cauchy	OLS (196)	LP (185)	LP (142)
Uniform	OLS (198)	LP (197)	LP (192)
Laplace	OLS (200)	LP (200)	LP (192)

Case 14 : Descending Variance, p=-0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (198)	LP (198)	LP (192)
Cauchy	OLS (194)	LP (178)	LP (138)
Uniform	OLS (200)	LP (193)	LP (187)
Laplace	OLS (199)	LP (199)	LP (188)

Case 15 : Variance of two greater than that of other three , p=-0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (199)	LP (198)	LP (194)
Cauchy	OLS (197)	LP (179)	LP (137)
Uniform	OLS (199)	LP (197)	LP (190)
Laplace	OLS (200)	LP (200)	LP (191)

Case 16 : Equal Variance, p=-0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (198)	LP (197)	LP (191)
Cauchy	OLS (190)	LP (180)	LP (147)
Uniform	OLS (197)	LP (193)	LP (187)
Laplace	OLS (198)	LP (196)	LP (188)

Case 17 : Descending Variance, p=-0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (197)	LP (194)	LP (187)
Cauchy	OLS (187)	LP (178)	LP (145)
Uniform	OLS (194)	LP (190)	LP (185)
Laplace	OLS (198)	LP (196)	LP (189)

Case 18 : Variance of two greater than that of other three , p=-0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (198)	LP (192)
Cauchy	OLS (196)	LP (189)	LP (149)
Uniform	OLS (199)	LP (196)	LP (190)

Laplace	OLS (200)	LP (200)	LP (193)

Case 19 : Equal Variance, p=-0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (199)	LP (198)	LP (192)
Cauchy	OLS (197)	LP (186)	LP (148)
Uniform	OLS (198)	LP (193)	LP (189)
Laplace	OLS (200)	LP (200)	LP (193)

Case 20 : Descending Variance, p=-0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (197)	LP (192)
Cauchy	OLS (195)	LP (183)	LP (144)
Uniform	OLS (196)	LP (193)	LP (188)
Laplace	OLS (200)	LP (199)	LP (191)

Case 21 : Variance of two greater than that of other three , p=-0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (193)
Cauchy	OLS (198)	LP (192)	LP (147)
Uniform	OLS (200)	LP (199)	LP (194)
Laplace	OLS (200)	LP (200)	LP (195)

(B) Normal Sample Size (T=30)

Case 1 : Equal Variance, p=0

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)*	LP (200)	LP (199)
Cauchy	OLS (200)	LP (199)	LP (153)
Uniform	OLS (200)	LP (200)	LP (200)
Laplace	OLS (200)	LP (200)	LP (198)

* The number of times the winning method ranks the first in the two hundred replicates. This applies to following analyses.

Case 2 : Descending Variance, p=0

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (200)
Cauchy	OLS (199)	LP (198)	LP (154)
Uniform	OLS (199)	LP (199)	LP (198)
Laplace	OLS (200)	LP (200)	LP (198)

Case 3 : Variance of two greater than that of other three, p=0

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (200)
Cauchy	OLS (200)	LP (198)	LP (153)
Uniform	OLS (200)	LP (200)	LP (200)
Laplace	OLS (200)	LP (200)	LP (197)

Case 4 : Equal Variance, p=0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (199)
Cauchy	OLS (200)	LP (198)	LP (149)

Uniform	OLS (200)	LP (200)	LP (199)
Laplace	OLS (200)	LP (200)	LP (197)

Case 5 : Descending Variance, p=0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (199)	LP (200)	LP (196)
Cauchy	OLS (200)	LP (195)	LP (149)
Uniform	OLS (200)	LP (200)	LP (196)
Laplace	OLS (200)	LP (200)	LP (195)

Case 6 : Variance of two greater than that of other three , p=0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (198)
Cauchy	OLS (199)	LP (197)	LP (151)
Uniform	OLS (200)	LP (200)	LP (197)
Laplace	OLS (200)	LP (200)	LP (198)

Case 7 : Equal Variance, p=0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (196)
Cauchy	OLS (199)	LP (196)	LP (149)
Uniform	OLS (200)	LP (200)	LP (195)
Laplace	OLS (200)	LP (200)	LP (193)

Case 8 : Descending Variance, p=0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (197)	LP (198)	LP (194)

Cauchy	OLS (196)	LP (192)	LP (136)
Uniform	OLS (199)	LP (198)	LP (197)
Laplace	OLS (199)	LP (197)	LP (192)

<u>Case 9</u> : Variance of two greater than that of other three , p=0.5

<u>Distribution</u>	<u>MSE Rank</u>	<u>MAD Rank</u>	<u>APE Rank</u>
Normal	OLS (198)	LP (199)	LP (195)
Cauchy	OLS (199)	LP (192)	LP (139)
Uniform	OLS (200)	LP (200)	LP (197)
Laplace	OLS (200)	LP (200)	LP (196)

<u>Case 10</u> : Equal Variance, p=0.8

<u>Distribution</u>	<u>MSE Rank</u>	<u>MAD Rank</u>	<u>APE Rank</u>
Normal	OLS (198)	LP (199)	LP (193)
Cauchy	OLS (197)	LP (190)	LP (147)
Uniform	OLS (198)	LP (196)	LP (191)
Laplace	OLS (199)	LP (197)	LP (195)

<u>Case 11</u> : Descending Variance, p=0.8

<u>Distribution</u>	<u>MSE Rank</u>	<u>MAD Rank</u>	<u>APE Rank</u>
Normal	OLS (178)	LP (177)	LP (165)
Cauchy	OLS (186)	LP (173)	LP (126)
Uniform	OLS (181)	LP (179)	LP (163)
Laplace	OLS (180)	LP (173)	LP (164)

<u>Case 12</u> : Variance of two greater than that of other three , p=0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (192)	LP (196)	LP (191)
Cauchy	OLS (194)	LP (184)	LP (136)
Uniform	OLS (195)	LP (197)	LP (185)
Laplace	OLS (193)	LP (192)	LP (184)

Case 13 : Equal Variance, p=-0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (199)
Cauchy	OLS (200)	LP (199)	LP (151)
Uniform	OLS (200)	LP (200)	LP (199)
Laplace	OLS (200)	LP (200)	LP (200)

Case 14 : Descending Variance, p=-0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (199)
Cauchy	OLS (200)	LP (200)	LP (149)
Uniform	OLS (200)	LP (200)	LP (199)
Laplace	OLS (200)	LP (200)	LP (200)

Case 15 : Variance of two greater than that of other three , p=-0.2

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (198)
Cauchy	OLS (200)	LP (200)	LP (150)
Uniform	OLS (200)	LP (200)	LP (199)
Laplace	OLS (200)	LP (200)	LP (200)

Case 16 : Equal Variance, p=-0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (193)
Cauchy	OLS (200)	LP (200)	LP (137)
Uniform	OLS (200)	LP (200)	LP (194)
Laplace	OLS (200)	LP (200)	LP (196)

Case 17 : Descending Variance, p=-0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (194)
Cauchy	OLS (200)	LP (200)	LP (139)
Uniform	OLS (200)	LP (200)	LP (194)
Laplace	OLS (200)	LP (200)	LP (195)

Case 18 : Variance of two greater than that of other three , p=-0.5

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (194)
Cauchy	OLS (200)	LP (200)	LP (138)
Uniform	OLS (200)	LP (200)	LP (193)
Laplace	OLS (200)	LP (200)	LP (195)

Case 19 : Equal Variance, p=-0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (194)
Cauchy	OLS (200)	LP (199)	LP (145)
Uniform	OLS (200)	LP (200)	LP (194)
Laplace	OLS (200)	LP (200)	LP (196)

Case 20 : Descending Variance, p=-0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (194)
Cauchy	OLS (200)	LP (199)	LP (146)
Uniform	OLS (200)	LP (200)	LP (193)
Laplace	OLS (200)	LP (200)	LP (197)

Case 21 : Variance of two greater than that of other three , p=-0.8

Distribution	MSE Rank	MAD Rank	APE Rank
Normal	OLS (200)	LP (200)	LP (194)
Cauchy	OLS (200)	LP (199)	LP (145)
Uniform	OLS (200)	LP (199)	LP (194)
Laplace	OLS (200)	LP (200)	LP (194)

BIBLIOGRAPHY

Books

Bolt Gordon J. (1988) Market and Sales Forecasting. 2nd ed. New York: Franklin Watts, Inc. 275-276

Churchill G A, Ford N M and Walker O C (1993) Sales Force Management. 4th ed. Homewood, Illinois: Richard D. Irwin Inc. 186-187, 204

Eby F H and O'Neill W J (1977) The Management of Sales Forecasting. Lexington, Massachusetts: D.C. Heath and Company, 1-4

Makridakis S and Wheelwright Steven C (1989) Forecasting Methods for Management. 5th ed. Singapore: John Wiley & Sons, Inc. 13-15

Wilson H J and Keating B (1990) Business Forecasting. Homewood, Illinois: Richard D. Irwin Inc. 330-344

Periodicals

Bates J M and Granger C W J (1969) "The Combination of Forecasts" Operations Research Quarterly. 20 (December), 451-68

Bunn D W (1989) "Editorial: Forecasting With More Than One Model" Journal of Forecasting. 8 (July-September), 161-166

------ (1975) "A Bayesian Approach to the Linear Combination of Forecasts" Operational Research Quarterly. 26 325-329

Clemen R T (1986) "Linear Constraints and the Efficiency of Combined Forecasts" Journal of Forecasting. 5 (1) 31-38

------ and Winkler R (1985) "Limits for the Prediction and Value of Information From Dependent Sources" Operations Research Quarterly. 33(March-April) 427-442

------ (1989) "Combining Forecasts: A Review and Annotated Bibliography" International Journal of Forecasting. 5 (4) 559-584

Dickinson J P (1973) "Some Statistical Results on the Combination of Forecasts" Operational Research Quarterly. 24 253-260

Granger C W J (1989) "Invited Review: Combining Forecasts-- Twenty Years Later" Journal of Forecasting. 8 (July-September) 167-173

------ and Newbold P (1973) "Some Comments on the Evaluation of Economic Forecasts" Applied Economics. 5 (March) 35-47

------ and Ramanathan R (1984) "Improving Methods of Combining Forecasts" Journal of Forecasting. 3 (April-June) 197-204

Moriarty M M (1990) "Boundary Value Models for the Combination of Forecasts" Journal of Marketing Research. 27 (November) 402-417

Mills T C and Stephenson M J (1985) "Forecasting Contemporaneous Aggregates and the Combination of Forecasts: The Case of the U.K. Monetary Aggregates" Journal of Forecasting. 4 273-284

Newbold P, Zumwalt K and Kannan S (1987) "Combining Forecasts to Improve Earnings Per Share Prediction: An Examination of Electric Utilities" International Journal of Forecasting. 3 (2) 229-238

Russell T D and Adam E E (1987) "An Empirical Evaluation of Alternative Forecasting Combinations" Management Science. 33 (October) 1267-1276

Sessions D N and Chatterjee S (1989) "The Combining of Forecasts Using Recursive Techniques With Non-stationary Weights" Journal of Forecasting. 8 (July-September) 239-251

Smith S A , McIntyre S H and Achabal D D (1994) "A Two-stage Sales Forecasting Procedures Using Discounted Least Squares" Journal of Market Research. 31 (February) 44-56

Weitz B A (1985) "Introduction to Special Issue on Competition in Marketing" Journal of Market Research. 22 (August) 229-236